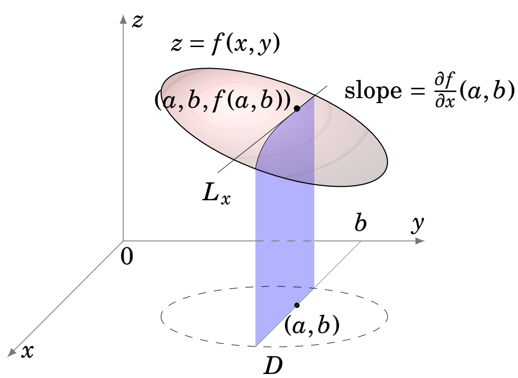
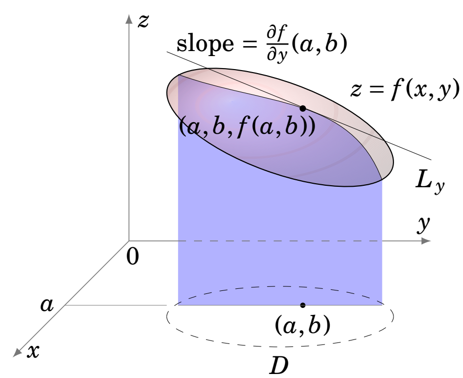
2.3 Tangent Plane to a Surface

In the previous section we mentioned that the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ can be thought of as the rate of change of a function $z = f(x,y)$ in the positive x and y directions, respectively. Recall that the derivative $\frac{dy}{dx}$ of a function $y = f(x)$ has a geometric meaning, namely as the slope of the tangent line to the graph of f at the point $(x, f(x))$ in $\mathbb{R}^2$. There is a similar geometric meaning to the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ of a function $z = f(x,y)$: given a point (a,b) in the domain D of $f(x,y)$, the trace of the surface described by $z = f(x,y)$ in the plane $y = b$ is a curve in $\mathbb{R}^3$ through the point $(a,b,f(a,b))$, and the slope of the tangent line L_x to that curve at that point is $\frac{\partial f}{\partial x}(a,b)$. Similarly, $\frac{\partial f}{\partial y}(a,b)$ is the slope of the tangent line L_y to the trace of the surface $z = f(x,y)$ in the plane $x = a$ (see Figure 2.3.1).



(a) Tangent line L_x in the plane $y = b$



(b) Tangent line L_y in the plane $x = a$

Figure 2.3.1 Partial derivatives as slopes

Since the derivative $\frac{dy}{dx}$ of a function $y = f(x)$ is used to find the tangent line to the graph of f (which is a curve in $\mathbb{R}^2$), you might expect that partial derivatives can be used to define a *tangent plane* to the graph of a surface $z = f(x,y)$. This indeed turns out to be the case. First, we need a definition of a tangent plane. The intuitive idea is that a tangent plane "just touches" a surface at a point. The formal definition mimics the intuitive notion of a tangent line to a curve.

Definition 2.4. Let $z = f(x,y)$ be the equation of a surface S in $\mathbb{R}^3$, and let $P = (a,b,c)$ be a point on S. Let T be a plane which contains the point P, and let $Q = (x,y,z)$ represent a generic point on the surface S. If the (acute) angle between the vector $\overrightarrow{PQ}$ and the plane T approaches zero as the point Q approaches P along the surface S, then we call T the **tangent plane** to S at P.

Note that since two lines in $\mathbb{R}^3$ determine a plane, then the two tangent lines to the surface $z = f(x,y)$ in the x and y directions described in Figure 2.3.1 are contained in the tangent plane at that point, *if the tangent plane exists at that point*. The existence of those two

tangent lines does not by itself guarantee the existence of the tangent plane. It is possible that if we take the trace of the surface in the plane $x - y = 0$ (which makes a 45° angle with the positive x-axis), the resulting curve in that plane may have a tangent line which is not in the plane determined by the other two tangent lines, or it may not have a tangent line at all at that point. Luckily, it turns out[4] that if $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist in a region around a point (a,b) and are continuous at (a,b) then the tangent plane to the surface $z = f(x,y)$ will exist at the point $(a,b,f(a,b))$. In this text, those conditions will always hold.

Suppose that we want an equation of the tangent plane T to the surface $z = f(x,y)$ at a point $(a,b,f(a,b))$. Let L_x and L_y be the tangent lines to the traces of the surface in the planes $y = b$ and $x = a$, respectively (as in Figure 2.3.2), and suppose that the conditions for T to exist do hold. Then the equation for T is

$$A(x-a) + B(y-b) + C(z - f(a,b)) = 0 \tag{2.4}$$

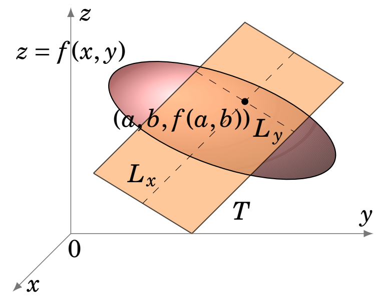


Figure 2.3.2 Tangent plane

where $\mathbf{n} = (A,B,C)$ is a normal vector to the plane T. Since T contains the lines L_x and L_y, then all we need are vectors $\mathbf{v}_x$ and $\mathbf{v}_y$ that are parallel to L_x and L_y, respectively, and then let $\mathbf{n} = \mathbf{v}_x \times \mathbf{v}_y$.

Since the slope of L_x is $\frac{\partial f}{\partial x}(a,b)$, then the vector $\mathbf{v}_x = (1,0,\frac{\partial f}{\partial x}(a,b))$ is parallel to L_x (since $\mathbf{v}_x$ lies in the xz-plane and lies in a line with slope $\frac{\frac{\partial f}{\partial x}(a,b)}{1} = \frac{\partial f}{\partial x}(a,b)$. See Figure 2.3.3). Similarly, the vector $\mathbf{v}_y = (0,1,\frac{\partial f}{\partial y}(a,b))$ is parallel to L_y. Hence, the vector

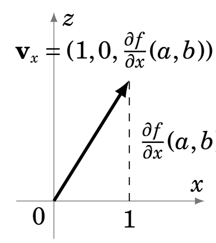


Figure 2.3.3

$$\mathbf{n} = \mathbf{v}_x \times \mathbf{v}_y = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & \frac{\partial f}{\partial x}(a,b) \\ 0 & 1 & \frac{\partial f}{\partial y}(a,b) \end{vmatrix} = -\frac{\partial f}{\partial x}(a,b)\,\mathbf{i} - \frac{\partial f}{\partial y}(a,b)\,\mathbf{j} + \mathbf{k}$$

is normal to the plane T. Thus the equation of T is

$$-\frac{\partial f}{\partial x}(a,b)(x-a) - \frac{\partial f}{\partial y}(a,b)(y-b) + z - f(a,b) = 0\,. \tag{2.5}$$

Multiplying both sides by -1, we have the following result:

The equation of the tangent plane to the surface $z = f(x,y)$ at the point $(a,b,f(a,b))$ is

$$\frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) - z + f(a,b) = 0 \tag{2.6}$$

[4]See TAYLOR and MANN, § 6.4.

Example 2.13. Find the equation of the tangent plane to the surface $z = x^2 + y^2$ at the point $(1,2,5)$.

Solution: For the function $f(x,y) = x^2 + y^2$, we have $\frac{\partial f}{\partial x} = 2x$ and $\frac{\partial f}{\partial y} = 2y$, so the equation of the tangent plane at the point $(1,2,5)$ is

$$2(1)(x-1) + 2(2)(y-2) - z + 5 = 0 \text{ , or}$$
$$2x + 4y - z - 5 = 0 \text{ .}$$

In a similar fashion, it can be shown that if a surface is defined implicitly by an equation of the form $F(x,y,z) = 0$, then the tangent plane to the surface at a point (a,b,c) is given by the equation

$$\frac{\partial F}{\partial x}(a,b,c)(x-a) + \frac{\partial F}{\partial y}(a,b,c)(y-b) + \frac{\partial F}{\partial z}(a,b,c)(z-c) = 0 \text{ .} \tag{2.7}$$

Note that formula (2.6) is the special case of formula (2.7) where $F(x,y,z) = f(x,y) - z$.

Example 2.14. Find the equation of the tangent plane to the surface $x^2 + y^2 + z^2 = 9$ at the point $(2,2,-1)$.

Solution: For the function $F(x,y,z) = x^2 + y^2 + z^2 - 9$, we have $\frac{\partial F}{\partial x} = 2x$, $\frac{\partial F}{\partial y} = 2y$, and $\frac{\partial F}{\partial z} = 2z$, so the equation of the tangent plane at $(2,2,-1)$ is

$$2(2)(x-2) + 2(2)(y-2) + 2(-1)(z+1) = 0 \text{ , or}$$
$$2x + 2y - z - 9 = 0 \text{ .}$$

Exercises

A

For Exercises 1-6, find the equation of the tangent plane to the surface $z = f(x,y)$ at the point P.

1. $f(x,y) = x^2 + y^3$, $P = (1,1,2)$

2. $f(x,y) = xy$, $P = (1,-1,-1)$

3. $f(x,y) = x^2 y$, $P = (-1,1,1)$

4. $f(x,y) = xe^y$, $P = (1,0,1)$

5. $f(x,y) = x + 2y$, $P = (2,1,4)$

6. $f(x,y) = \sqrt{x^2 + y^2}$, $P = (3,4,5)$

For Exercises 7-10, find the equation of the tangent plane to the given surface at the point P.

7. $\frac{x^2}{4} + \frac{y^2}{9} + \frac{z^2}{16} = 1$, $P = \left(1, 2, \frac{2\sqrt{11}}{3}\right)$

8. $x^2 + y^2 + z^2 = 9$, $P = (0,0,3)$

9. $x^2 + y^2 - z^2 = 0$, $P = (3,4,5)$

10. $x^2 + y^2 = 4$, $P = (\sqrt{3},1,0)$

2.4 Directional Derivatives and the Gradient

For a function $z = f(x,y)$, we learned that the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ represent the (instantaneous) rate of change of f in the positive x and y directions, respectively. What about other directions? It turns out that we can find the rate of change in *any* direction using a more general type of derivative called a *directional derivative*.

Definition 2.5. Let $f(x,y)$ be a real-valued function with domain D in $\mathbb{R}^2$, and let (a,b) be a point in D. Let $\mathbf{v}$ be a unit vector in $\mathbb{R}^2$. Then the **directional derivative of f at (a,b) in the direction of $\mathbf{v}$**, denoted by $D_{\mathbf{v}} f(a,b)$, is defined as

$$D_{\mathbf{v}} f(a,b) = \lim_{h\to 0} \frac{f((a,b)+h\mathbf{v}) - f(a,b)}{h} \tag{2.8}$$

Notice in the definition that we seem to be treating the point (a,b) as a vector, since we are adding the vector $h\mathbf{v}$ to it. But this is just the usual idea of identifying vectors with their terminal points, which the reader should be used to by now. If we were to write the vector $\mathbf{v}$ as $\mathbf{v} = (v_1, v_2)$, then

$$D_{\mathbf{v}} f(a,b) = \lim_{h\to 0} \frac{f(a+hv_1, b+hv_2) - f(a,b)}{h} \; . \tag{2.9}$$

From this we can immediately recognize that the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are special cases of the directional derivative with $\mathbf{v} = \mathbf{i} = (1,0)$ and $\mathbf{v} = \mathbf{j} = (0,1)$, respectively. That is, $\frac{\partial f}{\partial x} = D_{\mathbf{i}} f$ and $\frac{\partial f}{\partial y} = D_{\mathbf{j}} f$. Since there are many vectors with the same direction, we use a unit vector in the definition, as that represents a "standard" vector for a given direction.

If $f(x,y)$ has continuous partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ (which will always be the case in this text), then there is a simple formula for the directional derivative:

Theorem 2.2. Let $f(x,y)$ be a real-valued function with domain D in $\mathbb{R}^2$ such that the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist and are continuous in D. Let (a,b) be a point in D, and let $\mathbf{v} = (v_1, v_2)$ be a unit vector in $\mathbb{R}^2$. Then

$$D_{\mathbf{v}} f(a,b) = v_1 \frac{\partial f}{\partial x}(a,b) + v_2 \frac{\partial f}{\partial y}(a,b) \; . \tag{2.10}$$

Proof: Note that if $\mathbf{v} = \mathbf{i} = (1,0)$ then the above formula reduces to $D_{\mathbf{v}} f(a,b) = \frac{\partial f}{\partial x}(a,b)$, which we know is true since $D_{\mathbf{i}} f = \frac{\partial f}{\partial x}$, as we noted earlier. Similarly, for $\mathbf{v} = \mathbf{j} = (0,1)$ the formula reduces to $D_{\mathbf{v}} f(a,b) = \frac{\partial f}{\partial y}(a,b)$, which is true since $D_{\mathbf{j}} f = \frac{\partial f}{\partial y}$. So since $\mathbf{i} = (1,0)$ and $\mathbf{j} = (0,1)$ are the only unit vectors in $\mathbb{R}^2$ with a zero component, then we need only show the formula holds for unit vectors $\mathbf{v} = (v_1, v_2)$ with $v_1 \neq 0$ and $v_2 \neq 0$. So fix such a vector $\mathbf{v}$ and fix a number $h \neq 0$.

Then

$$f(a+hv_1,b+hv_2)-f(a,b)=f(a+hv_1,b+hv_2)-f(a+hv_1,b)+f(a+hv_1,b)-f(a,b)\ . \quad (2.11)$$

Since $h \neq 0$ and $v_2 \neq 0$, then $hv_2 \neq 0$ and thus any number c between b and $b+hv_2$ can be written as $c = b+\alpha hv_2$ for some number $0 < \alpha < 1$. So since the function $f(a+hv_1, y)$ is a real-valued function of y (since $a+hv_1$ is a fixed number), then the Mean Value Theorem from single-variable calculus can be applied to the function $g(y) = f(a+hv_1, y)$ on the interval $[b, b+hv_2]$ (or $[b+hv_2, b]$ if one of h or v_2 is negative) to find a number $0 < \alpha < 1$ such that

$$\frac{\partial f}{\partial y}(a+hv_1,b+\alpha hv_2)=g'(b+\alpha hv_2)=\frac{g(b+hv_2)-g(b)}{b+hv_2-b}=\frac{f(a+hv_1,b+hv_2)-f(a+hv_1,b)}{hv_2}$$

and so

$$f(a+hv_1,b+hv_2)-f(a+hv_1,b) = hv_2\frac{\partial f}{\partial y}(a+hv_1,b+\alpha hv_2)\ .$$

By a similar argument, there exists a number $0 < \beta < 1$ such that

$$f(a+hv_1,b)-f(a,b) = hv_1\frac{\partial f}{\partial x}(a+\beta hv_1,b)\ .$$

Thus, by equation (2.11), we have

$$\begin{aligned}\frac{f(a+hv_1,b+hv_2)-f(a,b)}{h} &= \frac{hv_2\frac{\partial f}{\partial y}(a+hv_1,b+\alpha hv_2)+hv_1\frac{\partial f}{\partial x}(a+\beta hv_1,b)}{h}\\ &= v_2\frac{\partial f}{\partial y}(a+hv_1,b+\alpha hv_2)+v_1\frac{\partial f}{\partial x}(a+\beta hv_1,b)\end{aligned}$$

so by formula (2.9) we have

$$\begin{aligned}D_{\mathbf{v}}f(a,b) &= \lim_{h\to 0}\frac{f(a+hv_1,b+hv_2)-f(a,b)}{h}\\ &= \lim_{h\to 0}\left[v_2\frac{\partial f}{\partial y}(a+hv_1,b+\alpha hv_2)+v_1\frac{\partial f}{\partial x}(a+\beta hv_1,b)\right]\\ &= v_2\frac{\partial f}{\partial y}(a,b)+v_1\frac{\partial f}{\partial x}(a,b) \quad \text{by the continuity of } \frac{\partial f}{\partial x} \text{ and } \frac{\partial f}{\partial y}, \text{ so}\\ D_{\mathbf{v}}f(a,b) &= v_1\frac{\partial f}{\partial x}(a,b)+v_2\frac{\partial f}{\partial y}(a,b)\end{aligned}$$

after reversing the order of summation. **QED**

Note that $D_{\mathbf{v}}f(a,b) = \mathbf{v}\cdot\left(\frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b)\right)$. The second vector has a special name:

Definition 2.6. For a real-valued function $f(x,y)$, the **gradient** of f, denoted by ∇f, is the vector

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \tag{2.12}$$

in $\mathbb{R}^2$. For a real-valued function $f(x,y,z)$, the gradient is the vector

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \tag{2.13}$$

in $\mathbb{R}^3$. The symbol ∇ is pronounced "del".[5]

Corollary 2.3. $D_{\mathbf{v}} f = \mathbf{v} \cdot \nabla f$

Example 2.15. Find the directional derivative of $f(x,y) = xy^2 + x^3y$ at the point $(1,2)$ in the direction of $\mathbf{v} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.

Solution: We see that $\nabla f = (y^2 + 3x^2y, 2xy + x^3)$, so

$$D_{\mathbf{v}} f(1,2) \;=\; \mathbf{v} \cdot \nabla f(1,2) \;=\; \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right) \cdot (2^2 + 3(1)^2(2), 2(1)(2) + 1^3) \;=\; \tfrac{15}{\sqrt{2}}$$

A real-valued function $z = f(x,y)$ whose partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist and are continuous is called *continuously differentiable*. Assume that $f(x,y)$ is such a function and that $\nabla f \neq \mathbf{0}$. Let c be a real number in the range of f and let $\mathbf{v}$ be a unit vector in $\mathbb{R}^2$ which is tangent to the level curve $f(x,y) = c$ (see Figure 2.4.1).

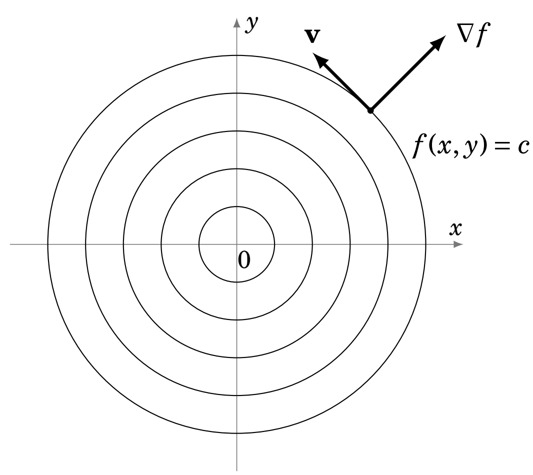


Figure 2.4.1

[5]Sometimes the notation grad(f) is used instead of ∇f.

The value of $f(x,y)$ is constant along a level curve, so since $\mathbf{v}$ is a tangent vector to this curve, then the rate of change of f in the direction of $\mathbf{v}$ is 0, i.e. $D_{\mathbf{v}}f = 0$. But we know that $D_{\mathbf{v}}f = \mathbf{v}\cdot\nabla f = \|\mathbf{v}\|\,\|\nabla f\|\cos\theta$, where θ is the angle between $\mathbf{v}$ and ∇f. So since $\|\mathbf{v}\| = 1$ then $D_{\mathbf{v}}f = \|\nabla f\|\cos\theta$. So since $\nabla f \neq \mathbf{0}$ then $D_{\mathbf{v}}f = 0 \Rightarrow \cos\theta = 0 \Rightarrow \theta = 90°$. In other words, $\nabla f \perp \mathbf{v}$, which means that ∇f is *normal* to the level curve.

In general, for any unit vector $\mathbf{v}$ in $\mathbb{R}^2$, we still have $D_{\mathbf{v}}f = \|\nabla f\|\cos\theta$, where θ is the angle between $\mathbf{v}$ and ∇f. At a fixed point (x,y) the length $\|\nabla f\|$ is fixed, and the value of $D_{\mathbf{v}}f$ then varies as θ varies. The largest value that $D_{\mathbf{v}}f$ can take is when $\cos\theta = 1$ ($\theta = 0°$), while the smallest value occurs when $\cos\theta = -1$ ($\theta = 180°$). In other words, the value of the function f increases the fastest in the direction of ∇f (since $\theta = 0°$ in that case), and the value of f decreases the fastest in the direction of $-\nabla f$ (since $\theta = 180°$ in that case). We have thus proved the following theorem:

Theorem 2.4. Let $f(x,y)$ be a continuously differentiable real-valued function, with $\nabla f \neq \mathbf{0}$. Then:

(a) The gradient ∇f is normal to any level curve $f(x,y) = c$.

(b) The value of $f(x,y)$ increases the fastest in the direction of ∇f.

(c) The value of $f(x,y)$ decreases the fastest in the direction of $-\nabla f$.

Example 2.16. In which direction does the function $f(x,y) = xy^2 + x^3y$ increase the fastest from the point $(1,2)$? In which direction does it decrease the fastest?

Solution: Since $\nabla f = (y^2 + 3x^2y, 2xy + x^3)$, then $\nabla f(1,2) = (10,5) \neq \mathbf{0}$. A unit vector in that direction is $\mathbf{v} = \frac{\nabla f}{\|\nabla f\|} = \left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$. Thus, f increases the fastest in the direction of $\left(\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$ and decreases the fastest in the direction of $\left(\frac{-2}{\sqrt{5}}, \frac{-1}{\sqrt{5}}\right)$.

Though we proved Theorem 2.4 for functions of two variables, a similar argument can be used to show that it also applies to functions of three or more variables. Likewise, the directional derivative in the three-dimensional case can also be defined by the formula $D_{\mathbf{v}}f = \mathbf{v}\cdot\nabla f$.

Example 2.17. The temperature T of a solid is given by the function $T(x,y,z) = e^{-x} + e^{-2y} + e^{4z}$, where x, y, z are space coordinates relative to the center of the solid. In which direction from the point $(1,1,1)$ will the temperature decrease the fastest?

Solution: Since $\nabla f = (-e^{-x}, -2e^{-2y}, 4e^{4z})$, then the temperature will decrease the fastest in the direction of $-\nabla f(1,1,1) = (e^{-1}, 2e^{-2}, -4e^{4})$.

Exercises

A

For Exercises 1-10, compute the gradient ∇f.

1. $f(x,y) = x^2 + y^2 - 1$

2. $f(x,y) = \dfrac{1}{x^2 + y^2}$

3. $f(x,y) = \sqrt{x^2 + y^2 + 4}$

4. $f(x,y) = x^2 e^y$

5. $f(x,y) = \ln(xy)$

6. $f(x,y) = 2x + 5y$

7. $f(x,y,z) = \sin(xyz)$

8. $f(x,y,z) = x^2 e^{yz}$

9. $f(x,y,z) = x^2 + y^2 + z^2$

10. $f(x,y,z) = \sqrt{x^2 + y^2 + z^2}$

For Exercises 11-14, find the directional derivative of f at the point P in the direction of $\mathbf{v} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.

11. $f(x,y) = x^2 + y^2 - 1,\ P = (1,1)$

12. $f(x,y) = \dfrac{1}{x^2 + y^2},\ P = (1,1)$

13. $f(x,y) = \sqrt{x^2 + y^2 + 4},\ P = (1,1)$

14. $f(x,y) = x^2 e^y,\ P = (1,1)$

For Exercises 15-16, find the directional derivative of f at the point P in the direction of $\mathbf{v} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right)$.

15. $f(x,y,z) = \sin(xyz),\ P = (1,1,1)$

16. $f(x,y,z) = x^2 e^{yz},\ P = (1,1,1)$

17. Repeat Example 2.16 at the point $(2,3)$.

18. Repeat Example 2.17 at the point $(3,1,2)$.

B

For Exercises 19-26, let $f(x,y)$ and $g(x,y)$ be continuously differentiable real-valued functions, let c be a constant, and let $\mathbf{v}$ be a unit vector in $\mathbb{R}^2$. Show that:

19. $\nabla(cf) = c\nabla f$

20. $\nabla(f+g) = \nabla f + \nabla g$

21. $\nabla(fg) = f\nabla g + g\nabla f$

22. $\nabla(f/g) = \dfrac{g\nabla f - f\nabla g}{g^2}$ if $g(x,y) \neq 0$

23. $D_{-\mathbf{v}} f = -D_{\mathbf{v}} f$

24. $D_{\mathbf{v}}(cf) = c D_{\mathbf{v}} f$

25. $D_{\mathbf{v}}(f+g) = D_{\mathbf{v}} f + D_{\mathbf{v}} g$

26. $D_{\mathbf{v}}(fg) = f D_{\mathbf{v}} g + g D_{\mathbf{v}} f$

27. The function $r(x,y) = \sqrt{x^2 + y^2}$ is the length of the position vector $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$ for each point (x,y) in $\mathbb{R}^2$. Show that $\nabla r = \dfrac{1}{r}\mathbf{r}$ when $(x,y) \neq (0,0)$, and that $\nabla(r^2) = 2\mathbf{r}$.

2.5 Maxima and Minima

The gradient can be used to find *extreme points* of real-valued functions of several variables, that is, points where the function has a *local maximum* or *local minimum*. We will consider only functions of two variables; functions of three or more variables require methods using linear algebra.

Definition 2.7. Let $f(x,y)$ be a real-valued function, and let (a,b) be a point in the domain of f. We say that f has a **local maximum** at (a,b) if $f(x,y) \le f(a,b)$ for all (x,y) inside some disk of positive radius centered at (a,b), i.e. there is some sufficiently small $r > 0$ such that $f(x,y) \le f(a,b)$ for all (x,y) for which $(x-a)^2 + (y-b)^2 < r^2$.

Likewise, we say that f has a **local minimum** at (a,b) if $f(x,y) \ge f(a,b)$ for all (x,y) inside some disk of positive radius centered at (a,b).

If $f(x,y) \le f(a,b)$ for all (x,y) in the domain of f, then f has a **global maximum** at (a,b). If $f(x,y) \ge f(a,b)$ for all (x,y) in the domain of f, then f has a **global minimum** at (a,b).

Suppose that (a,b) is a local maximum point for $f(x,y)$, and that the first-order partial derivatives of f exist at (a,b). We know that $f(a,b)$ is the largest value of $f(x,y)$ as (x,y) goes in all directions from the point (a,b), in some sufficiently small disk centered at (a,b). In particular, $f(a,b)$ is the largest value of f in the x direction (around the point (a,b)), that is, the single-variable function $g(x) = f(x,b)$ has a local maximum at $x = a$. So we know that $g'(a) = 0$. Since $g'(x) = \frac{\partial f}{\partial x}(x,b)$, then $\frac{\partial f}{\partial x}(a,b) = 0$. Similarly, $f(a,b)$ is the largest value of f near (a,b) in the y direction and so $\frac{\partial f}{\partial y}(a,b) = 0$. We thus have the following theorem:

Theorem 2.5. Let $f(x,y)$ be a real-valued function such that both $\frac{\partial f}{\partial x}(a,b)$ and $\frac{\partial f}{\partial y}(a,b)$ exist. Then a necessary condition for $f(x,y)$ to have a local maximum or minimum at (a,b) is that $\nabla f(a,b) = \mathbf{0}$.

Note: Theorem 2.5 can be extended to apply to functions of three or more variables.

A point (a,b) where $\nabla f(a,b) = \mathbf{0}$ is called a **critical point** for the function $f(x,y)$. So given a function $f(x,y)$, to find the critical points of f you have to solve the equations $\frac{\partial f}{\partial x}(x,y) = 0$ and $\frac{\partial f}{\partial y}(x,y) = 0$ simultaneously for (x,y). Similar to the single-variable case, the *necessary* condition that $\nabla f(a,b) = \mathbf{0}$ is not always *sufficient* to guarantee that a critical point is a local maximum or minimum.

Example 2.18. The function $f(x,y) = xy$ has a critical point at $(0,0)$: $\frac{\partial f}{\partial x} = y = 0 \Rightarrow y = 0$, and $\frac{\partial f}{\partial y} = x = 0 \Rightarrow x = 0$, so $(0,0)$ is the only critical point. But clearly f does not have a local maximum or minimum at $(0,0)$ since any disk around $(0,0)$ contains points (x,y) where the values of x and y have the same sign (so that $f(x,y) = xy > 0 = f(0,0)$) and different signs (so that $f(x,y) = xy < 0 = f(0,0)$). In fact, along the path $y = x$ in $\mathbb{R}^2$, $f(x,y) = x^2$, which has a

local minimum at $(0,0)$, while along the path $y = -x$ we have $f(x, y) = -x^2$, which has a local maximum at $(0,0)$. So $(0,0)$ is an example of a *saddle point*, i.e. it is a local maximum in one direction and a local minimum in another direction. The graph of $f(x, y)$ is shown in Figure 2.5.1, which is a hyperbolic paraboloid.

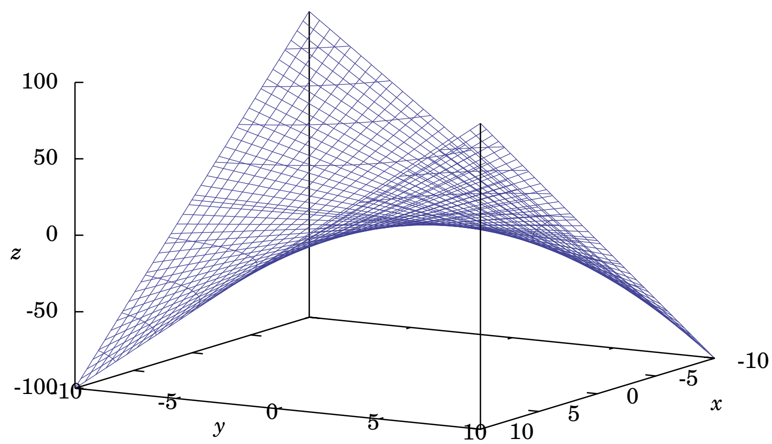


Figure 2.5.1 $f(x, y) = xy$, saddle point at $(0,0)$

The following theorem gives sufficient conditions for a critical point to be a local maximum or minimum of a *smooth function* (i.e. a function whose partial derivatives of all orders exist and are continuous), which we will not prove here.[6]

Theorem 2.6. Let $f(x, y)$ be a smooth real-valued function, with a critical point at (a, b) (i.e. $\nabla f(a, b) = \mathbf{0}$). Define

$$D = \frac{\partial^2 f}{\partial x^2}(a,b)\,\frac{\partial^2 f}{\partial y^2}(a,b) - \left(\frac{\partial^2 f}{\partial y\,\partial x}(a,b)\right)^2$$

Then

(a) if $D > 0$ and $\frac{\partial^2 f}{\partial x^2}(a,b) > 0$, then f has a local minimum at (a, b)

(b) if $D > 0$ and $\frac{\partial^2 f}{\partial x^2}(a,b) < 0$, then f has a local maximum at (a, b)

(c) if $D < 0$, then f has neither a local minimum nor a local maximum at (a, b)

(d) if $D = 0$, then the test fails.

[6] See TAYLOR and MANN, § 7.6.

If condition (c) holds, then (a,b) is a *saddle point*. Note that the assumption that $f(x,y)$ is smooth means that

$$D = \begin{vmatrix} \dfrac{\partial^2 f}{\partial x^2}(a,b) & \dfrac{\partial^2 f}{\partial y\,\partial x}(a,b) \\ \dfrac{\partial^2 f}{\partial x\,\partial y}(a,b) & \dfrac{\partial^2 f}{\partial y^2}(a,b) \end{vmatrix}$$

since $\frac{\partial^2 f}{\partial y\,\partial x} = \frac{\partial^2 f}{\partial x\,\partial y}$. Also, if $D > 0$ then $\frac{\partial^2 f}{\partial x^2}(a,b)\frac{\partial^2 f}{\partial y^2}(a,b) = D + \left(\frac{\partial^2 f}{\partial y\,\partial x}(a,b)\right)^2 > 0$, and so $\frac{\partial^2 f}{\partial x^2}(a,b)$ and $\frac{\partial^2 f}{\partial y^2}(a,b)$ have the same sign. This means that in parts (a) and (b) of the theorem one can replace $\frac{\partial^2 f}{\partial x^2}(a,b)$ by $\frac{\partial^2 f}{\partial y^2}(a,b)$ if desired.

Example 2.19. Find all local maxima and minima of $f(x,y) = x^2 + xy + y^2 - 3x$.
Solution: First find the critical points, i.e. where $\nabla f = \mathbf{0}$. Since

$$\frac{\partial f}{\partial x} = 2x + y - 3 \quad \text{and} \quad \frac{\partial f}{\partial y} = x + 2y$$

then the critical points (x,y) are the common solutions of the equations

$$\begin{aligned} 2x + y - 3 &= 0 \\ x + 2y \phantom{{}-3} &= 0 \end{aligned}$$

which has the unique solution $(x,y) = (2,-1)$. So $(2,-1)$ is the only critical point.

To use Theorem 2.6, we need the second-order partial derivatives:

$$\frac{\partial^2 f}{\partial x^2} = 2\,, \quad \frac{\partial^2 f}{\partial y^2} = 2\,, \quad \frac{\partial^2 f}{\partial y\,\partial x} = 1$$

and so

$$D \;=\; \frac{\partial^2 f}{\partial x^2}(2,-1)\frac{\partial^2 f}{\partial y^2}(2,-1) - \left(\frac{\partial^2 f}{\partial y\,\partial x}(2,-1)\right)^2 \;=\; (2)(2) - 1^2 \;=\; 3 \;>\; 0$$

and $\frac{\partial^2 f}{\partial x^2}(2,-1) = 2 > 0$. Thus, $(2,-1)$ is a local minimum.

Example 2.20. Find all local maxima and minima of $f(x,y) = xy - x^3 - y^2$.
Solution: First find the critical points, i.e. where $\nabla f = \mathbf{0}$. Since

$$\frac{\partial f}{\partial x} = y - 3x^2 \quad \text{and} \quad \frac{\partial f}{\partial y} = x - 2y$$

then the critical points (x, y) are the common solutions of the equations

$$\begin{aligned} y - 3x^2 &= 0 \\ x - 2y &= 0 \end{aligned}$$

The first equation yields $y = 3x^2$, substituting that into the second equation yields $x - 6x^2 = 0$, which has the solutions $x = 0$ and $x = \frac{1}{6}$. So $x = 0 \Rightarrow y = 3(0) = 0$ and $x = \frac{1}{6} \Rightarrow y = 3\left(\frac{1}{6}\right)^2 = \frac{1}{12}$. So the critical points are $(x, y) = (0,0)$ and $(x, y) = \left(\frac{1}{6}, \frac{1}{12}\right)$.

To use Theorem 2.6, we need the second-order partial derivatives:

$$\frac{\partial^2 f}{\partial x^2} = -6x\ , \quad \frac{\partial^2 f}{\partial y^2} = -2\ , \quad \frac{\partial^2 f}{\partial y\,\partial x} = 1$$

So

$$D \;=\; \frac{\partial^2 f}{\partial x^2}(0,0)\,\frac{\partial^2 f}{\partial y^2}(0,0) - \left(\frac{\partial^2 f}{\partial y\,\partial x}(0,0)\right)^2 \;=\; (-6(0))(-2) - 1^2 \;=\; -1 \;<\; 0$$

and thus $(0,0)$ is a saddle point. Also,

$$D \;=\; \frac{\partial^2 f}{\partial x^2}\left(\tfrac{1}{6}, \tfrac{1}{12}\right)\frac{\partial^2 f}{\partial y^2}\left(\tfrac{1}{6}, \tfrac{1}{12}\right) - \left(\frac{\partial^2 f}{\partial y\,\partial x}\left(\tfrac{1}{6}, \tfrac{1}{12}\right)\right)^2 \;=\; (-6(\frac{1}{6}))(-2) - 1^2 \;=\; 1 \;>\; 0$$

and $\frac{\partial^2 f}{\partial x^2}\left(\frac{1}{6}, \frac{1}{12}\right) = -1 < 0$. Thus, $\left(\frac{1}{6}, \frac{1}{12}\right)$ is a local maximum.

Example 2.21. Find all local maxima and minima of $f(x, y) = (x-2)^4 + (x-2y)^2$.

Solution: First find the critical points, i.e. where $\nabla f = \mathbf{0}$. Since

$$\frac{\partial f}{\partial x} = 4(x-2)^3 + 2(x-2y) \quad \text{and} \quad \frac{\partial f}{\partial y} = -4(x-2y)$$

then the critical points (x, y) are the common solutions of the equations

$$\begin{aligned} 4(x-2)^3 + 2(x-2y) &= 0 \\ -4(x-2y) &= 0 \end{aligned}$$

The second equation yields $x = 2y$, substituting that into the first equation yields $4(2y-2)^3 = 0$, which has the solution $y = 1$, and so $x = 2(1) = 2$. Thus, $(2,1)$ is the only critical point.

To use Theorem 2.6, we need the second-order partial derivatives:

$$\frac{\partial^2 f}{\partial x^2} = 12(x-2)^2 + 2\ , \quad \frac{\partial^2 f}{\partial y^2} = 8\ , \quad \frac{\partial^2 f}{\partial y\,\partial x} = -4$$

So

$$D = \frac{\partial^2 f}{\partial x^2}(2,1)\,\frac{\partial^2 f}{\partial y^2}(2,1) - \left(\frac{\partial^2 f}{\partial y\,\partial x}(2,1)\right)^2 = (2)(8)-(-4)^2 = 0$$

and so the test fails. What can be done in this situation? Sometimes it is possible to examine the function to see directly the nature of a critical point. In our case, we see that $f(x,y) \ge 0$ for all (x,y), since $f(x,y)$ is the sum of fourth and second powers of numbers and hence must be nonnegative. But we also see that $f(2,1) = 0$. Thus $f(x,y) \ge 0 = f(2,1)$ for all (x,y), and hence $(2,1)$ is in fact a *global* minimum for f.

Example 2.22. Find all local maxima and minima of $f(x,y) = (x^2+y^2)e^{-(x^2+y^2)}$.

Solution: First find the critical points, i.e. where $\nabla f = \mathbf{0}$. Since

$$\begin{aligned}\frac{\partial f}{\partial x} &= 2x(1-(x^2+y^2))e^{-(x^2+y^2)}\\ \frac{\partial f}{\partial y} &= 2y(1-(x^2+y^2))e^{-(x^2+y^2)}\end{aligned}$$

then the critical points are $(0,0)$ and all points (x,y) on the unit circle $x^2+y^2=1$.

To use Theorem 2.6, we need the second-order partial derivatives:

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= 2[1-(x^2+y^2)-2x^2-2x^2(1-(x^2+y^2))]e^{-(x^2+y^2)}\\ \frac{\partial^2 f}{\partial y^2} &= 2[1-(x^2+y^2)-2y^2-2y^2(1-(x^2+y^2))]e^{-(x^2+y^2)}\\ \frac{\partial^2 f}{\partial y\,\partial x} &= -4xy[2-(x^2+y^2)]e^{-(x^2+y^2)}\end{aligned}$$

At $(0,0)$, we have $D = 4 > 0$ and $\frac{\partial^2 f}{\partial x^2}(0,0) = 2 > 0$, so $(0,0)$ is a local minimum. However, for points (x,y) on the unit circle $x^2+y^2=1$, we have

$$D = (-4x^2e^{-1})(-4y^2e^{-1}) - (-4xye^{-1})^2 = 0$$

and so the test fails. If we look at the graph of $f(x,y)$, as shown in Figure 2.5.2, it looks like we might have a local maximum for (x,y) on the unit circle $x^2+y^2=1$. If we switch to using polar coordinates (r,θ) instead of (x,y) in $\mathbb{R}^2$, where $r^2 = x^2+y^2$, then we see that we can write $f(x,y)$ as a function $g(r)$ of the variable r alone: $g(r) = r^2e^{-r^2}$. Then $g'(r) = 2r(1-r^2)e^{-r^2}$, so it has a critical point at $r=1$, and we can check that $g''(1) = -4e^{-1} < 0$, so the Second Derivative Test from single-variable calculus says that $r=1$ is a local maximum. But $r=1$ corresponds to the unit circle $x^2+y^2=1$. Thus, the points (x,y) on the unit circle $x^2+y^2=1$ are local maximum points for f.

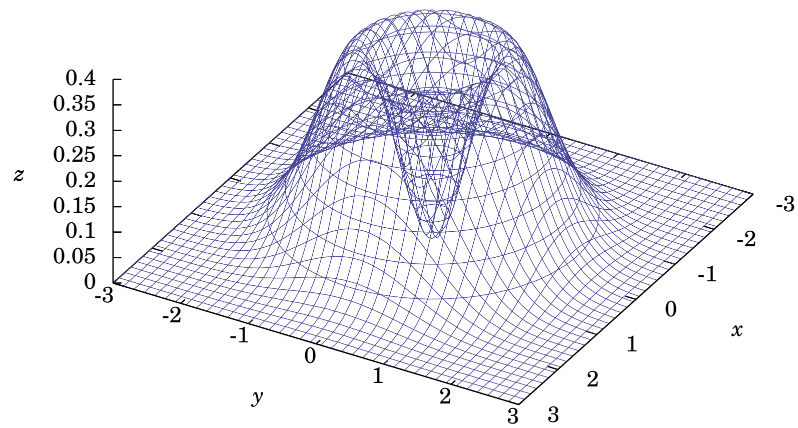


Figure 2.5.2 $f(x,y)=(x^2+y^2)e^{-(x^2+y^2)}$

Exercises

A

For Exercises 1-10, find all local maxima and minima of the function $f(x,y)$.

1. $f(x,y)=x^3-3x+y^2$

2. $f(x,y)=x^3-12x+y^2+8y$

3. $f(x,y)=x^3-3x+y^3-3y$

4. $f(x,y)=x^3+3x^2+y^3-3y^2$

5. $f(x,y)=2x^3+6xy+3y^2$

6. $f(x,y)=2x^3-6xy+y^2$

7. $f(x,y)=\sqrt{x^2+y^2}$

8. $f(x,y)=x+2y$

9. $f(x,y)=4x^2-4xy+2y^2+10x-6y$

10. $f(x,y)=-4x^2+4xy-2y^2+16x-12y$

B

11. For a rectangular solid of volume 1000 cubic meters, find the dimensions that will minimize the surface area. (*Hint: Use the volume condition to write the surface area as a function of just two variables.*)

12. Prove that if (a,b) is a local maximum or local minimum point for a smooth function $f(x,y)$, then the tangent plane to the surface $z=f(x,y)$ at the point $(a,b,f(a,b))$ is parallel to the xy-plane. (*Hint: Use Theorem 2.5.*)

C

13. Find three positive numbers x, y, z whose sum is 10 such that x^2y^2z is a maximum.